DOCUMENT RESUME

ED 210 825 EA 014 309

AUTHOR Thiemann, Francis C.
TITLE Semiotic Analysis.
PUB DATE [79]
NOTE 20p.

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Data Analysis; Definitions; Intellectual History; Logic; Matrices; Organizational Theories; *Semiotics

ABSTRACT

Semiotic analysis is a method of analyzing signs (e.g., words) to reduce non-numeric data to their component parts without losing essential meanings. Semiotics dates back to Aristotle's analysis of language; it was much advanced by nineteenth-century analyses of style and logic and by Whitehead and Russell's description in this century of the role of language in logic. Recent writers have applied semiotics to sociological data and content analysis. The method of semiotic analysis comprises, first, definition reduction and, second, identification of interrelationships. In definition reduction, a frequency dictionary of all words used is created, from which synonyms are drawn and categories created. The categories are then combined in a symmetrical matrix and logical decision rules are used to identify their interrelationships. An example of semiotic analysis illustrates how to apply it to a set of descriptive propositions concerning a new administrator in a school organization. Semiotic analysis may be used in case or policy studies or in any area where many words are used to describe phenomena. Organizational theorists can use semiotic analysis to reduce inconsistencies, identify central issues or parameters, and analyze observational notes. (Author/EW)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED210825

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it.
X Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Francis C. Thiemann

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

SEMIOTIC ANALYSIS

DR. FRANCIS C. THIEMANN
UNIVERSITY OF LOUISVILLE, BELKNAP CAMPUS
LOUISVILLE, KENTUCKY

EA 014 309

DRAFT

SEMIOTIC ANALYSIS

Abstract

Semiotic Analysis (SA) is a rational process of reducing large quantities of alpha string data[1] without losing the essential meaning of the individual items. A series of decision rules determine how the data are organized, categorized, and interrelated.

In general, semiotics is concerned with signs and their application, especially in language as developed and systematized within Scientific Empericism. Specifically, semiotics or the theory of signs is a generic term which includes three distinct branches: pragmatics, semantics, and syntactics. Pragmatics focuses on the relation between signs and those who produce or receive and under stand them. Semantics examines the relation between signs and what they refer to (designata or denotata). It also contains the theory of truth (semantical definition) and the theory of logical deduction. Lastly, syntatics is concerned with the formal relations among signs which includes logical syntax (theoretical language, formal calculi, and formal, symbolic, mathematical logic).

It is to be understood then, that Semiotic Analysis (SA) incorporates the concepts and methods of these disciplines in separating alpha data into its component parts or constituent elements.

I have attempted in this paper to lay out the answers to three specific questions: 1) What have been the major evolutionary stages in alpha data analysis that have lead to Semiotic Analysis? 2) How does Semiotic Analysis (SA) work (a short example has been included); and 3) What are the kinds of problems for which SA can be used?

In attempting to carry out this task, I may be like the small boy who upon receiving a book on penguins from his grandmother was asked how he liked it. He replied, "I already know more about penguins than I care to." But, for those of you who are avid readers of newspaper fillins, I hope to add to your excitement and to your repertoire.

WHAT HAVE BEEN THE MAJOR EVOLUTIONARY STAGES IN ALPHA DATA ANALYSIS THAT HAVE LED TO SEMIOTIC ANALYSIS?

Like so many other fields of study, Aristotle was the first to explore, under the name of analytics, the techniques of language analysis. In Prior Analytics Aristotle examines the truth logic of the syllogism. Although concerned with the content of statements (their truth value), Aristotle was primarily interested in the structure of language that caused some[2] interrelated statements to be considered valid. In Posteriori Analytics, he turned his attention to identifying the necessary conditions required to establish scientific or verifiable knowledge.

For the next many centuries through dormant and active periods, these Aristotelian concepts were examined and extended. Not until the mid 1800's with his Mathematical Analysis of Logic (1847) and Laws of Thought (1854), did George Boole significantly modify alpha data analysis methodology. It was with his work that language analysys entered the modern era.

On a slightly different wave in 1867, Lewis Campbell, the famous English Plationist, suggested that the language of any careful writer may undergo a discernible evolution. So, if the evolutionary pattern were detected, then the chronology of undated works could be ordered. Thirty years later (1897), Wincenty Lutoslawski published The Origin and Growth of Plato's Logic. It was in this work that Lutoslawski presented his well developed method of studying the evolution of style which he called styleometrics. In brief, the styleometric method is one in which words are counted and classed according

to one of four "peculiarities" depending upon their frequence and importance. Works of similiar length were compared and as Lutoslawski states in his "Law of Stylistic Affinity" (p. 152).

> Of two works of the same author and of the same size, that is nearer in time to a third, which shares with it the greater number of stylistic peculiarities, provided that their different importance is taken into account, and that the number of observed peculiarities is sufficient to determine the stylistic character of all three work.

With this method of language analysis, Lutoslawski was able to determine the order, heretofore unknown, in which six of Plato's major works were written.

Although advancement were being made in alpha data analysis, the most rapid growth was in numeric data analysis (mathematics). Arguments waged back and forth between the need of verification and the creative intuition. Positivism continued to gain ground and there were not a few who held that underlying the structure of the Universe was a mathematical formula which if discovered would solve all man's riddles. The first puncture of the Positivist bubble came in the early 1900's with the 3 volume publication of the Principia Mathematica. In this work, Whitehead and Russell convincingly show that the construction of Indo-European languages is the basis of both the logic and the mathematics of Western civilization. They go on to prove (logically) that all logical relations between ideas and concrete reality can be reduced to conjunction (and) and negation (it is false that). Having this impetus semanicist furthered the argument by showing that by reflecting on language, which is the

symbolization of a referent that may or may not have real existence, we can sharpen our awareness of the similarities that exist between things (discrimination), since things are necessarily similar on the basis of their existence alone, if nothing more.

As these new ideas gained ground mathematics were retreating from accentuating the value of empirical abstraction and objectivity. The contemporary view, among mathematicians, is that the propositions of mathematics have nothing to say about empirical reality. Even applied geometry, the most reality based area of mathematics which supposedly organizes physical measurement, does not receive from geometicians an interpretation in which its propositions become unqualifiedly experimental or empirical. This is in no way intended to deny the validity of mathematical manipulation, it simply states, that since mathematics deals with the intellectual rather than substantial reality, its a priori definitions have an internal consistency and precision not found in a posteriori based analytics. What then is said for logic? Argument forms are valid or in-valid, premises are true or false; the validity of the argument does not convey truth to the conclusion when the premises are false--this is also true for mathematics. Brooks (1970) and Von Dusseldorp (1971) concerned themselves with this problem in the social sciences where the mathematical formula is often too rigorous for sloppy data.

While these opposing truths continued to maintain their bastions, however, the slowly evolving concepts of language analysis were beginning to be studied by a growing number of researchers. Zetterberg in 1954 published his first edition of On Theory and Verification in Sociology. Chapter 4 "On Propositions in Sociology" and Chapter 5 "on the Ordering of Sociological Proposition" were his attempts to

add the concepts of truth logic verification to the discipline of Sociology. Albeit primitive, from the standpoint of symbolic logic, Zetterberg does make a contribution with his definitional reduction in axiomatic format. Other sociologist soon attached themselves to the axiomatic concepts and (all too soon!) they imposed on it mathematical rigor. And eventually, rigormortis.

During this same period, the 1950's, Laswell, Lerner and Pool published The Comparative Study of Symbols (1952). The basic concepts of this work relied heavily on some of the same ideas advanced by Campbell and Lutoslawski. Holsti (1969) provides an excellent review and present stage of development in his Content Analysis for the Social Sciences and the Humanities. Given this brief and very incomplete description of the history of alpha analysis, I will now address the second question:

HOW DOES SEMIOTIC ANALYSIS WORK?

The entire Semiotic Analysis process may be conducted by a "man-alone"[3] or a "man-machine"[4] operation. When the volume of propositional data (if, then statements) is small-300 more or less items--a "man-alone" operation is recommended. With large volumes of interview transaction, observational notes or many more than 300 propositions the computere becomes essential.

Definition reduction is the first step of the semiotic process. In man-machine operations, a computer is used to develop a dictionary of all the individual words in a given data collection. When the data collected are from interviews or abservational notes, the verbatim transcript is fed directly into the computer. If, however, the collection is a set of statements or propositions abstracted from a

variety of sources or by a number of individuals, it may be necessary (or administratively convenient) to rewrite the items into a standard form. From a data bank, thus formed, a dictionary containing only the words given the computer is created. Words may be listed either in alphabetical, frequency order or in frequency, alphabetical order. In the former, the dictionary is built as an ordinary dictionary, while the frequency, alphabetical dictionary list alphabetical the most often used words first.

From such dictionaries an analyst may request a variety of word relations in accordance with his underlying theory or predetermined decision rules. Examples of some word relationship are discussed in Holsti (1969) as tag programs in which "...every entry word is defined with one or more 'tag' representing categories in the investigator's theory," (1969;156). Depending upon the a priori category, already known. While the a priori alternative may be used with a semiotic analysis dictionary, semiotic analysis also allows categories to be formed directly from the particular data collection--a posteriori. The advantages of a posteriori--derived categories have been well known since Aristotle and, therefore, will not be discussed here. One of the major advantages of Semiotic Analysis over other methods of Content Analysis (See Holsti 1969, 195-221) is establishing categories or sets and subsets inductively - a posteriori.

Following the older concepts of content analysis, semiotic analysis dictionaries always calculate the frequency of work occurrence. The frequencies provide some quick observational clues as to which words may be combined (two words with similar frequency counts have a higher probability of being juxtaposed than two with markedly different frequencies), but, more importantly, the frequency dictionary provides

a gestalt of illative importance to the analyst.

From the dictionaries, the analyst may also identify those phrases or families of words that may have synonyms which, in turn, may be used as substitutions for words in the original collection. The immediate effect of the synonym substitution is a reduction in the size of the original collection without a concomitant semantic reduction. Synonym substitution and volume reduction facilitates the creation of categories since key concepts and words are made more apparent; as are the the interrelationships between them.

When the a posteriori categories have been isolated from the collection, the analyst can then retrieve from the computer all those items that contribute to each set in the category. In some cases one item may appear in two or more sets between categories or within a category.

As indicated before, definition reduction can be accomplished by a man-alone operation. In much the same manner as the man-computer operation, man-alone attempts to isolate the key words and concepts by reading and sorting individual statements that have been previously typed on 3 x 5 cards. (Underlining key words or phrases can facilitate sorting.) Duplicate cards may be made for those statements containing two or more key words or concepts. While the units employed in the man-alone operation are grosser--sentences instead of words--and the sorting into the sets and subsets of major categories requires many more steps, the end result is virtually the same for similarly trained analysts.

In the second stage of the process, the items collected in each category (set and subset) are then combined in a symmetrical matrix. One of several decision rules could be made to identify their interrelationship. If the decision rule uses one of the connectives of formal logic, a deductive argument could be formed where the validity of the arguments formed and the truth value of the propositions could lead to a true conclusion (or new knowledge--Zetterberg). Since the conclusion implies the premises of the argument, the data have been reduced. Further reduction can occur by developing additional arguments using the conclusions as premises, and so on.

If other decision rule relationships are desired such as importance, sequence, and centrality, premises may be created by direct inference, arguments formed, and collection volume reduced.

I would now like to take you through a short example of a semiotic analysis.

Example 5

In a study of succession a number of hypothetical propositions were abstracted from the literature and initially categorized according to key actors such as predecessor, successor, outside successor, and suboordinate. The set of propositions concerned with the outside successor in a new role position are presented below to illustrate definition and logical reduction.

Step 1: List the propositions.

The outsider creates a new informal social circle.

The outsider is a stranger, unacquainted with the social realities of the situation.

The outsider can initially subscribe to the position of objectivity and impartiality.

The outsider does not know the informal social structure of the organization.

The outsider is unhampered by the informal social structure.

The outsider will have a difficult time being accepted.

The outsider is isolated from the informal sources of information.

The outsider is given considerable latitude in his operation.

The outsider is given full support of the board during the first year.

The outsider is at first over-accepted and over-idealized.

Step 2: Create a word frequency dictionary.

Word	Frequency of Occurrence	Word	Frequency of Occurrence
a	3	know	1
accepted	2	latitude	1
and	2	new	1
at	1	not	1
being	1	objectivity	1
board	1	of	5
by	1	operation	1
can	1	organization	1
circle	1	outsider	10
considerable	1	over	2
creates	1	position	1
difficult	1	realities	1
does	1	situation	1
during	1	social	4
from	1	sources	1
first	2	stranger	1

Word	Frequency of Occurrence	Word	Frequency of Occurrence
full	1	structure	2
given	2	subscribe	1
his	1	support	1
have	1	the	19
idealized	1	to	1
impartiality	1	time	1
in	1	unacquainted	1
informal	4	unhampered	1
information	1	will	1
initially	1	with	1
is	6	year	1
isolated	1		

In this dictionary, three words, "informal," "social," and "structure," have been used more than once and have more subsuming significance than the other words listed. The other multiple-use words are discounted because they either identify this particular set of propositions ("outsider") or are articles, prepositions, adverbs or verbs.

Step 3: Isolate the propositions containing the words "informal," "social," and "structure," and determine how may propositions can be subsumed.

The outsider creates a new informal social circle.

The outsider is a stranger, unacquainted with the social realities of the situation.

The outsider does not know the informal social structure of the organization.

The outsider is unhampered by the informal social structure.

The outsider is isolated from the informal sources of information.

Step 4: Write subsuming statement(s). In this case, the five statements can be represented by the single statement:

The outsider is a stranger.

Note: It is important that the meaning of the statements subsumed not be lost and that the statement "the outsider is a stranger" does not include or convey more than the statements from which it was drawn. Since the subsumed statements set the parameters of the subsuming statement, any subsequent use of the subsuming statement should be limited by these parameters.

Step 5: Replace the subsumed statements in the original data collection with the subsuming statement. The list of statements becomes reduced to six statements:

The outsider is a stranger. (Contained in this statement are the five subsumed statements.)

The outsider can initially subscribe to a position of objectivity and impartiality.

The outsider will have a difficult time being accepted.

The outsider is given considerable latitude in his operation.

The outsider is given full support of the board during the first year.

The outsider is at first over-accepted and over-idealized.

Step 6: Obtain a computer print out which places each statement in a hypothetical (if, then) relationship with each of the other statements. (Each statement should assume the positions of both antecedent and consequence.) If the volume of data is small, this step can be done by hand.

If the outsider is a stranger, then the outsider can initially subscribe to the position of objectivity and impartiality.

If the outsider is a stranger, then the outsider is given considerable latitude in his operation.

If the outsider is a stranger, then,. . . .

Step 7: Read each statement to determine if the meaning of the antecedent and consequence are the same or if the consequence further defines the antecedent. In both cases, the consequence can be subsumed and added to the original list of subsumed statements.

If the antecedent and consequence are not the same, an acceptance or rejection judgement is made. Devise a code for the list of propositions created in Step 6 to note acceptance/rejection decisions. For instance, a check (✓) can indicate that a relationship is accepted, a zero (0) that it is rejected.

Note: In reading the statements it is sometimes helpful to substitute the pronoun "he" for the term "the outsider." The reason for this is that without great care one can slip into accepting all the relationships by definition alone.

One of the statements, "the outsider will have a difficult time being accepted," was judged to be an additional parameter definind the subsuming statement, "the outsider is a stranger," and was therefore subsumed.

Step 8: Assign letter symbols to the final list of statements.

A The outsider is a stranger.

B The outsider can initially subscribe to the position of objectivity and impartiality.

C The outsider is given considerable latitude in his operation.

D The outsider is given full support of the board during the first year.

E The outsider is at first over-accepted and over-idealized.

Step 9: Place the five statements into a symmetrical matrix using the appropriate symbols to identify each statement and to note acceptance or rejection.

	A	B	C	D	E
A		A ⊃ B	A ⊃ C	A ⊃ D	A ⊃ E
B	.			B ⊃ D	.
C	.			C ⊃ D	.
D	D ⊃ A	D ⊃ B	D ⊃ C		D ⊃ E
E	.	E ⊃ B	E ⊃ C	E ⊃ D	

As shown in the chart, when a hypothetical relationship is accepted, a symbolic representation is placed in the appropriate cell--A⊃B, for example. If the lationship if reversible (accepted when read either way, for example, C⊃D and D⊃C), the symbols are placed in both cells. If one relationship is accepted and the other rejected, a dot (.) is placed in the rejected cell, and if neither relationship is accepted, the cells are X'ed (X) out.

Step 10: Create a valid argument which uses or implies all the relationships noted in the matrix.

$$\begin{array}{l} A \supset E \\ E \supset B \\ B \supset D \\ \therefore D \supset C \\ \hline \therefore A \supset C \end{array}$$

In ordinary language, this argument reads:

> If the outsider is a stranger, then he is at first over-accepted and over-idealized.
>
> If he is at first over-accepted and over-idealized, then he can initially subscribe to the position of objectivity amd impartiality.
>
> If he can initially subscribe to the position of objectivity and impartiality, then he is given full support of the board during the first year.
>
> If he is given full support of the board during the first year, then he is given considerable latitude in his operation.

Therefore, if the outsider is a stranger, then he is given considerable latitude in his operation.

By mediate or immediate inference, each filled cell of the matrix is contained in the argument.

Step 11: Following the logical reduction of each of the sets and subsets in the total collection, a matrix of conclusions is constructed and interrelationships are determined from which additional conclusions are drawn. In this way, a few conclusions can represent a large number of statements, and essential ideas can be isolated.

Recognize that this example represents a very small portion of much larger study. In reality the first part of semiotic analysis -- Definition Reduction would contain some 600 propositions that when separated into antecedent and consequence would represent 1200 items. The example of Logical Reduction represents part of one large metrics for the outside successor.

FOR WHAT KINDS OF PROBLEMS CAN SEMIOTIC ANALYSIS BE USED?

--case studies
--life crisis studies
--policy studies
--goal identification
--career competency
--or any area where large amounts of words are used to describe phenomena

Since this is a section on Organizational Thory, I will focus on my attention here. In the past, theory has been constructed by researchers, who have read widely in their field, have observed and experimented, or have past through the experiences themselves. They have then reflected on these vacarious or real experiences; abstracted what appeared to be critical, and project the possibility of interrelated events occuring again. The problem with such an approach is, of course, dependent upon the mind being able to abstract all the critical elements, to recall all the interrelated ones on demand, and most importantly, to interrelate them in developing a conclusion.

One of my old mentors, Arthur Symullian, once said at a symbolic logic conference, that most people could keep in mind and repeat in order three items when interrelated in a semetrical metric. Few could handle five and he was the only man he knew who could do seven--but not eight. That means given seven items, he could start at the top left hand corner of the metric and from his memory give the 42 possible combinations down to the second to the last cell in the bottom right hand corner. If this is true, and I do not doubt it, then is very unlikely, that with the number of variables, contingencies and exceptions with which most administrative or organizational researcher must deal in developing partial theories, that much real theory building has been done. That is, it is doubtful that any partial theory thus developed would not be riddled with internal inconsistencies.

This <u>is</u> an empirical question which could be tested by subjecting a theoretical work to semiotic analysis. For example, James G. Miller's, <u>Living System</u> perports to be a general theory which covers the total range of living systems from cell to society. Miller's hypotheses are clearly laid out as compound and complex sentences. They deal with everything from chemical and biological relations to decentralized decision making and societal boundaries. In the 765 pages of Living Systems, the probability of inconsistencies and contradictions is high, . . Therefore, SA could be used to reduce the errors that are inherent with volume.

On another tack, however, one might assume that Miller has covered general system as well as any man alive. Semiotic Analysis could be used, then, to isolate those areas where gaps occur, where central issued are concealed, where parameters can be established.

Another example, of this same problem and a little closer to home is Harry Wolcott's The Man in the Principal's Office: An Ethnography. Wolcott tells us:

> Prior to organizaing and sorting the information recorded on hundreds of pages of journal entries, the research process underwent a long gestation period for reading, discussing, mulling, and weighing alternative ways to excerpt and catalog the notes. Before converting the stream-of-consciousness notes, a tentative set of categories was developed from which the sections of the complete enthnography were finally organized. The categories guided the selection of relevant excerpts from the notes, and the notes in turn provided an ultimate test of the comprehensiveness of the categories. (Harry Wolcott, The Man in the Principal's Office: An Ethnography, 1973, p. 11).

It has not been uncommon for social science researcher to employ the methodology Wolcott describes, and it is almost with pride that they refer to their intuitive approach. They are, however, extremely defensive if questions of validity are raised. As one must do with the circular argument Wolcott presents: "The categories guided the selection of revelant excerpts from the notes, and the notes in turn provided an ultimate test of the comprehensiveness of the categories." How much better would it be to use the valuable observational notes in tact, to allow the categories to fall out naturally (a posteriori) and not to depend on the mind to do what a computer can do better. Believe me, I like what Harry did, it is fun to read and informative. But I think with more appropriate methodology and equal writing ability, Harry or someone else could do better.

FOOTNOTES

[1] Alpha string data differs from numeric string data in that words are the basic elements rather than numbers. The unit of analysis in alpha data is the meaning of sentences, statements, or propositions.

[2] Out of the 256 possible combinations of mood and figure only 24 are valid argument forms.

BIBLIOGRAPHY

Ross, W. D. (Trans.) THE WORKS OF ARISTOTLE. (London: Oxford University Press, 1952); McKeon, R. (Ed.), THE BASIC WORKS OF ARISTOTLE. (New York: Random House, 1941).

Boole, G. THE MATHEMATICAL ANALYSIS OF LOGIC. (New York: Philosophical Library, 1948).

Boole, G. AN INVESTIGATION OF THE LAWS OF THOUGHT. (New York: Dover Publications, 1951).

Lutoslawski, W. THE ORIGIN AND GROWTH OF PLATO'S LOGIC. (London: Longmans, Green, and Co., 1897).

Whitehead, A. N. and B. Russell. PRINCIPIA MATHEMATICS. (Cambridge, Eng.: The University Press, 1927-35).

Brookes, J. E. "Operational Research in Educational Administration" EDUCATION AND URBAN SOCIETY. 3 November, 1977.

Van Duggeldorp, P.A., et al. EDUCATIONAL DECISION-MAKING THROUGH OPERATIONS RESEARCH. (Boston: Allyn and Bacon, 1971).

Zetterberg, H. ON THEORY AND VERIFICATION IN SOCIOLOGY. (New Jersey: The Bedminster Press, 1965).

Laswell, H. D., et al. THE COMPARATIVE STUDY OF SYMBOLS. (Stanford: Stanford University Press, 1952).

Holsti, O. R. CONTENT ANALYSIS FOR THE SOCIAL SCIENCES AND HUMANITIES. Reading, Mass.: Addison-Wesley Publishing Co., 1969).

Thiemann, F. C. A PARTIAL THEORY OF EXECUTIVE SUCCESSION. (Unpublished Doctoral Dissentation, University of Oregon, 1968).

_____. ENVIRONMENTS AND PARADIGMS. (Edmonton, Alberta, Canada: University of Alberta, 1972).

Thiemann, F. C. and C. L. Borkosky. ARIOLE: A PLANNING GUIDE. (Eugene, Ore.: Center for Educational Policy and Management, 1974).

Miller, J. G. LIVING SYSTEM. (New York: McGraw-Hill, 1978).

Wolcott, H. THE MAN IN THE PRINICPAL'S OFFICE: AN ENTHNOGRAPHY. (New York: Holt, Rinehart and Winston, Inc., 1973).